

LISTING OF CLAIMS

1. (Currently Amended) A method for identifying groups of co-regulated and co-expressed genes, comprising:

defining a clustering criterion of data, relating to gene expression that varies with time and/or with the changing of environmental conditions, which is stored in a table;

in function of said clustering criterion, identifying in sub-tables groups of genes that satisfy said clustering criterion;

defining a number of logic filtering criteria of the data of said table;

for each logic filtering criterion, generating a corresponding filtered sub-table containing data of genes having expression values which satisfy said logic filtering criterion;

establishing pair combinations of said sub-tables by clustering and filtering the data of said table with said filtering criteria and by said clustering;

calculating characteristic parameters of the data associated to the groups of genes of each pair combination;

generating for each pair combination a characteristic value in function of the characteristic parameters of the groups of genes by a decision algorithm based on soft computing techniques;

identifying the groups of genes associated with pair combinations whose characteristic value is greater than a certain pre-established threshold as being members of a network of genes ~~likely to be~~ involved in a particular cellular process and discarding pair combinations of groups of genes whose characteristic value is smaller than said threshold; and

outputting the identified group of genes having characteristic values greater than ~~that~~ the threshold in a selected data format.

2. (Canceled).

3. (Previously Presented) The method of claim 1, wherein said decision algorithm is a fuzzy logic algorithm having antecedents and consequents which are defined in function of said parameters.

4. (Previously Presented) The method of claim 1, wherein said parameters are chosen from the group comprising numerical parameters tied to gene expression levels, parameters having a semantic biological meaning, and mixed parameters expressing at the same time a numerical relationship and a semantic meaning.

5. (Currently Amended) The method of claim 4,
wherein said numeric parameters are chosen from a group consisting of:
absolute values of linear correlation coefficients among data associated to pairs of genes;
absolute values of quadratic correlation coefficients among data associated to pairs of genes;
percentage of genes of the pair combination having a ~~the~~ final value of gene expression of which is greater than a ~~the~~ respective value of initial gene expression;
percentage of genes of the pair combination having a ~~the~~ final value of gene expression of which is smaller than a ~~the~~ respective value of initial gene expression;
percentage of genes having ~~the~~ values of gene expression of which have a same increasing or decreasing time evolution; and
percentage of genes that have a maximum value of gene expression in a same condition; and
wherein said parameters having a semantic biological meaning are chosen from a group consisting of:
percentage of genes that have ontologies in common; and
percentage of genes that have functional domains in common.

6. (Previously Presented) The method of claim 1, further comprising discarding combinations among sub-tables constituted by a number of genes smaller than a certain pre-established number, wherein genes that are comprised in both combined sub-tables are considered only once in connection with discarding.

7. (Previously Presented) The method of claim 1, wherein said clustering criteria are based on algorithms chosen in a set comprising: agglomerative hierarchic algorithms, non hierarchic Kmeans algorithms, hierarchic sequential Kmeans, non-hierarchic SOM and not exclusive Fuzzy Clustering.

8. (Currently Amended) The method of claim 5, comprising
calculating correlation coefficients of all pairs of gene sequences of the pair combination;
subdividing an ~~the~~ interval from 0 to 1 in five sub-intervals of equal length and assigning to each of said sub-intervals a respective quantized value of correlation;
calculating the percentage of correlation coefficients belonging to each sub-interval;
defining for each combination an overall coefficient of linear correlation obtained as arithmetic mean of the quantized values associated to the sub-intervals containing a number of coefficient greater than 50%.

9. (Withdrawn – Rejoinder Requested) The method of claim 5, comprising
calculating coefficients of quadratic correlation of all pairs of gene sequences of a same combination;
defining for each combination a global coefficient of quadratic correlation obtained as an arithmetic mean of said correlation values.

10. (Withdrawn – Rejoinder Requested – Currently Amended) The method of claim 5, comprising

calculating ~~a the~~ percentage of gene sequences ~~sequence~~ of the pair combination with a final value of gene expression greater than the initial value of gene expression;

defining a coefficient of global variation of the value of gene expression, comprised between 0 and 1, corresponding to said percentage.

11. (Withdrawn – Rejoinder Requested – Currently Amended) The method of claim 5, comprising

calculating ~~a the~~ percentage of gene sequences of the pair combination with an increasing time evolution;

defining a coefficient relative to the time evolution of the gene expression comprised between 0 and 1 corresponding to said percentage.

12. (Withdrawn – Rejoinder Requested – Currently Amended) The method of claim 5, comprising

calculating ~~a the~~ percentage of gene sequences of the pair combination with a value of gene expression greater than a pre-established threshold in correspondence of a same instant;

defining a coefficient of presence of maximum excursion of the level of gene expression in correspondence of the same instant, comprised between 0 and 1, corresponding to said percentage.

13. (Currently Amended) An identification system of groups of co-expressed and co-regulated genes, comprising:

a pre-processing sub-system input with data of a table relative to gene expressions variable with time and/or different environmental conditions, the pre-processing sub-system generating sub-tables of data in groups of genes that satisfy a pre-established clustering criterion;

a processing sub-system of data of said sub-tables, the processing sub-system considering all possible pairs of generated sub-tables and generating signals, for each pair of sub-tables, representing characteristic parameters of data associated to genes of that pair of sub-tables which express correlation among and between the included genes;

an intelligent sub-system input with said signals representative of characteristic parameters, generating for each pair of sub-tables a characteristic value determined as a function of the characteristic parameters and outputting data identifying groups of genes from each pair of sub-tables whose characteristic value is greater than a certain pre-established threshold as being members of a network of genes ~~likely to be~~ involved in a particular cellular process.

14. (Original) The system of claim 13, wherein said intelligent sub-system is a neural fuzzy logic sub-system, trained off-line.

15. (Currently Amended) A method for identifying groups of co-regulated and co-expressed genes, comprising:

receiving a table of data relating to evolution of gene expression with time and/or with changing environmental conditions for a plurality of genes;

applying a clustering algorithm to the table of data so as to identify clusters in the form of sub-tables comprising groups of genes that satisfy certain clustering criterion;

establishing all possible pair combinations of said clusters;

for each cluster pair combination, calculating a characteristic value for the cluster pair combination as a function of a plurality of characteristic parameters determined for each cluster pair which express a level of correlation which exists among and between the genes included in that cluster pair combination;

identifying the genes associated with cluster pair combinations whose characteristic value is greater than a certain pre-established threshold as being members of a network of genes likely to be involved in a particular cellular process; and

outputting the identified genes within the network in a selected data format.

16. (Currently Amended) A method for identifying groups of co-regulated and co-expressed genes, comprising:

receiving a table of data relating to evolution of gene expression with time and/or with changing environmental conditions for a plurality of genes;

applying a clustering algorithm to the table of data so as to identify clusters in the form of sub-tables comprising groups of genes that satisfy certain clustering criterion;

applying a filtering algorithm to the table of data so as to identify filter data in the form of sub-tables comprising groups of genes that satisfy certain filtering criteria;

establishing all possible pair combinations of said clusters, all possible pair combinations of filter data, and all possible pair combinations of clusters and filter data;

for each pair combination, calculating a characteristic value for the pair combination as a function of a plurality of characteristic parameters determined for each pair combination which express a level of correlation which exists among and between the genes included in that pair combination;

identifying the genes associated with pair combinations whose characteristic value is greater than a certain pre-established threshold as being members of a network of genes ~~likely to~~ be involved in a particular cellular process; and

outputting the identified genes within the network in a selected data format.

17. (Currently Amended) A method for the identification of groups of co-expressed and co-regulated genes, comprising:

receiving data in table format relative to gene expressions which are variable with time and/or different environmental conditions;

generating sub-tables of data in groups of genes that satisfy a pre-established clustering criterion;

identifying all possible pairs of generated sub-tables;

generating signals, for each pair of sub-tables, representing characteristic parameters of data associated to genes of that pair of sub-tables which express correlation among and between the included genes;

processing said signals representative of characteristic parameters to generate for each pair of sub-tables a characteristic value determined as a function of the characteristic parameters;

identifying groups of genes from each pair of sub-tables whose characteristic value is greater than a certain pre-established threshold;

outputting data in a selected format including the identified groups of genes as being members of a network of genes ~~likely to be~~ involved in a particular cellular process.